# Rapid formation of non-native contacts during the folding of HPr revealed by real-time photo-CIDNP NMR and stopped-flow fluorescence experiments

Canet, D; Lyon, CE; Scheek, RM; Robillard, GT; Dobson, CM; Hore, PJ; van Nuland, NAJ; Lyon, Charles E.; Dobson, Christopher M.; Hore, Peter J.

*Published in:*
Journal of Molecular Biology

*DOI:*
10.1016/S0022-2836(03)00507-2



*Publication date:*
2003

doi:10.1016/S0022-2836(03)00507-2 J. Mol. Biol. (2003) 330, 397–407

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

# Rapid Formation of Non-native Contacts During the Folding of HPr Revealed by Real-time Photo-CIDNP NMR and Stopped-flow Fluorescence Experiments

**Denis Canet[1]†, Charles E. Lyon[1]†, Ruud M. Scheek[2] George T. Robillard[2], Christopher M. Dobson[1], Peter J. Hore[1] and Nico A. J. van Nuland[3]***

[1]*Oxford Centre for Molecular Sciences and Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road Oxford OX1 3QZ, UK*

[2]*GBB, University of Groningen Nijenborgh 4, 9747 AG Groningen, The Netherlands*

[3]*Department of NMR Spectroscopy, Bijvoet Centre for Biomolecular Research University of Utrecht Padualaan 8, 3584 CH Utrecht The Netherlands*

**Corresponding author*

We report the combined use of real-time photo-CIDNP NMR and stopped-flow fluorescence techniques to study the kinetic refolding of a set of mutants of a small globular protein, HPr, in which each of the four phenylalanine residues has in turn been replaced by a tryptophan residue. The results indicate that after refolding is initiated, the protein collapses around at least three, and possibly all four, of the side-chains of these residues, as (i) the observation of transient histidine photo-CIDNP signals during refolding of three of the mutants (F2W, F29W, and F48W) indicates a strong decrease in tryptophan accessibility to the flavin dye; (ii) iodide quenching experiments show that the quenching of the fluorescence of F48W is less efficient for the species formed during the dead-time of the stopped-flow experiment than for the fully native state; and (iii) kinetic fluorescence anisotropy measurements show that the tryptophan side-chain of F48W has lower mobility in the dead-time intermediate state than in both the fully denatured and fully native states. The hydrophobic collapse observed for HPr during the early stages of its folding appears to act primarily to bury hydrophobic residues. This process may be important in preventing the protein from aggregating prior to the acquisition of native-like structure in which hydrophobic residues are exposed in order to play their role in the function of the protein. The phenylalanine residue at position 48 is likely to be of particular interest in this regard as it is involved in the binding to enzymes I and II that mediates the transfer of a phosphoryl group between the two enzymes.



*Keywords:* protein folding; non-native contacts; real-time NMR; fluorescence; photo-CIDNP

## Introduction

The manner in which a protein molecule achieves its unique native state during folding is one of the most dramatic examples of the way in which evolutionary pressure has generated molecules with properties that are able to generate diversity and selectivity in biology.[1–3] As well as resulting in the efficient production of functional proteins, the folding process is now recognized as being coupled to a variety of other biological processes including protein trafficking and the regulation of the cell cycle.[4] In addition, proteins have evolved to remain soluble in the complex milieu of the crowded biological environment, and to have the ability to interact specifically with their natural partners but not with other molecules. It is increasingly evident that the loss of control and regulation of cellular processes is the origin of a range of pathological conditions. These include amyloidosis and other protein degradation

† These two authors contributed equally to this work.

Present addresses: D. Canet, GeneProt Inc., 2 Pre-de-la-Fontaine, 1217 Meyrin/GE, Switzerland; C. M. Dobson, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

Abbreviations used: WT, wild-type; GdnDCL, deuterated guanidine hydrochloride; photo-CIDNP, photo-chemically induced dynamic nuclear polarization.

E-mail address of the corresponding author: nuland@nmr.chem.uu.nl

diseases, including many conditions associated with old age such as Alzheimer's and Parkinson's diseases.[3,5]

The manner by which a protein is able to achieve its functional state is beginning to emerge from experimental investigations of the structural changes occurring during protein folding. A more detailed understanding of such processes will shed light not only on how structure is encoded by the fold, but on the principles of the design of biological molecules that enable proteins to fold efficiently rather than to aggregate to species that may prevent completion of the folding process. Although *in vivo* molecular chaperones and other factors play an important role in controlling aggregation and other aspects of folding, the intrinsic properties of the sequence are likely to be the dominant factor that enables a protein to fold efficiently. It is therefore of particular importance to understand the detailed mechanism of protein folding *in vitro*, and to identify the fundamental determinants of the folding process.

The combination of kinetic NMR experiments with photo-CIDNP spectroscopy, in which enhanced nuclear spin polarization is generated by a laser flash, represents a powerful approach to characterizing in a direct manner important aspects of the structure of a protein during the folding process.[6] The strength of the photo-CIDNP approach lies not only in its time resolution but also in the reduction in spectral crowding arising from the polarization of only a limited number of residues (tyrosine, tryptophan and, under some conditions, histidine and methionine). Moreover, polarization occurs only when these residues are accessible to the photo-excited dye molecules in the solution, enabling the environment of individual residues to be probed directly in real time as folding takes place. This approach has been used to study the folding of lysozyme[7] and α-lactalbumin in the absence[8] and presence of $Ca^{2+}$ ions.[9] The results of these experiments suggest that for both proteins a relatively disorganised collapsed state is initially formed, and that reorganization then occurs to generate the native structure. Experiments of this type are therefore extremely valuable in studying details of folding processes, although care is needed to ensure that intensity effects are correctly interpreted in such complex systems where more than one CIDNP-active residue is present.[10] This consideration prompted us to use the photo-CIDNP approach to study the folding of single-tryptophan mutants of a protein such that the analysis of the CIDNP effects can be carried out for individual tryptophan residues located at selected positions within the overall fold.

The histidine-containing phosphocarrier protein HPr from *Escherichia coli* is a small 85-residue protein that is well suited to such an approach. The wild-type (WT) protein contains no tyrosine or tryptophan residues, but four phenylalanine residues that are particularly appropriate as mutation sites for replacement by tryptophan. Figure 1 shows a representation of the solution structure of the protein, showing that three α-helices are packed against a four-stranded antiparallel β-sheet.[11] Figure 1 also shows the location in the structure of the four phenylalanine residues. The replacement of each of the phenylalanine residues in turn with single tryptophan residues, resulting in four single-tryptophan mutants F2W, F22W, F29W, and F48W, does not result in substantial structural changes, but the presence of the bulkier side-chain leads to localized rearrangements around the mutated site.[31] The tryptophan side-chains are in similar environments to the corresponding phenylalanine side-chains in the WT structure, with that of residue 48 completely solvent accessible, those of residues 2 and 29 partially buried, and that of residue 22 completely buried in the core of the protein, inaccessible to the solvent. Replacement of the phenylalanine residues associated with the hydrophobic core of the protein, particularly those at positions 22 and 29, results in lower stability, most likely as a result of the introduction of the larger sized tryptophan side-chain that disrupts slightly the optimal packing of the core.[31] Figure 1 shows in addition the two histidine residues in HPr; His15 in the active-site is fully exposed to the solvent, while the His76 is less accessible. These residues were not mutated in the present study but act as reporters of the overall exposure of the tryptophan residues as we show below.

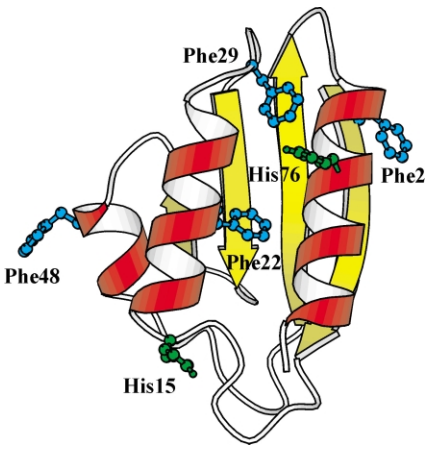


**Figure 1**. Molscript representation[30] of WT HPr. The side-chains of the four phenylalanine residues that were mutated in the present study, and the two histidine residues, are shown. Each of the four mutants studied here contains a single tryptophan residue, two histidine residues, and no tyrosine residues.

Here, we show that the aromatic tryptophan side-chains become inaccessible to solvent during the initial stages of folding of HPr, in some cases by forming non-native contacts that have to be broken in order to form the final native state structure. Results from fluorescence anisotropy and quenching studies are fully consistent with these photo-CIDNP observations and provide evidence for a rigid and relatively non-specific collapse of the

protein driven by the energetic advantage of sequestering the hydrophobic residues from solvent. At a later stage in folding the acquisition of native-like structure in the protein results in the exposure of certain hydrophobic groups, e.g. the aromatic residue at position 48, where they are able to perform their role in enabling HPr to transfer a phosphoryl group from the cytoplasmic to the membrane-bound enzyme (enzyme I and II, respectively) that results in the phosphorylation and transport of carbohydrates across the cell membranes of bacteria.[12]

## Results

### Photo-CIDNP of native and denatured states of mutants and wild-type HPr

Figure 2 shows the photo-CIDNP NMR spectra of WT HPr and the four single-tryptophan mutants in their native and fully unfolded (6 M guanidine hydrochloride, GdnDCl) states. The possibility of creating polarisation is limited to the single tryptophan and the two histidine residues, and only occurs when these residues are accessible to the photo-excited flavin molecules. When more than one side-chain is able to react with excited flavin dye (e.g. histidine and tryptophan in native F48W HPr), the CIDNP intensities of both amino acids depend critically on their ability to compete for the low concentration of triplet flavin radicals.[10] Excited flavin molecules can be quenched by (i) reaction with molecular oxygen, (ii) reaction with exposed histidine residues, or by (iii) reaction with the exposed tryptophan residue. The magnitudes of the second order rate constants for the reaction of histidine and tryptophan residues with excited flavin dye, $k_{His}$ and $k_{Trp}$, respectively have been estimated for the free amino acids by both transient absorption[13–15] and binary photo-CIDNP competition experiments.[10] In both cases it was found that $k_{Trp}$ is very much larger than $k_{His}$ (40:1 in experiments carried out under conditions similar to those utilized in the present study[10]).

The native state spectra of all the mutant proteins show relative intensities for the tryptophan and histidine resonances in the CIDNP spectra that are consistent with the relative positions of the phenylalanine residues in the three-dimensional structure of the WT protein (see Figure 1). These observations are in agreement with previous results from activity measurements, NMR experiments, and steady-state and time-resolved fluorescence measurements where it was found that the replacement of the phenylalanine residues with tryptophan at the various sites in HPr did not result in any substantial structural changes, but only localized rearrangements around the mutated sites.[31] In the denatured states of all four mutant proteins, the tryptophan residue in each case is fully exposed, consistent with a random coil model for an unfolded state, and competes so

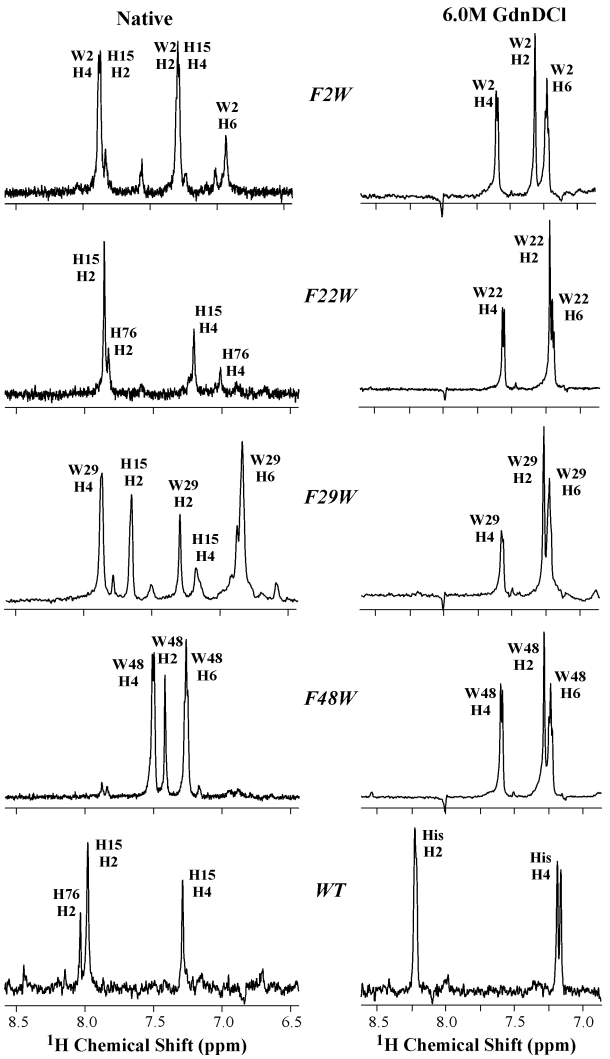


**Figure 2**. Photo-CIDNP spectra of the four single-tryptophan mutants of HPr and of the WT protein. Only the aromatic regions are shown. Polarization is limited to tryptophan and histidine residues, and can occur only when these residues are accessible to the photo-excited flavin molecules in solution, enabling the environment of individual residues to be probed directly. Spectra were recorded either in the absence (left panel) or presence (right panel) of 6.0 M GdnDCl. Differences in signal-to-noise are due to differences in protein concentrations used.

efficiently for the flavin dye so that no histidine signals are observed. However, when the tryptophan side-chain is partially (native F29W) or completely buried (native F22W) in the hydrophobic core of the protein, the histidine residues can compete successfully for the flavin dye and their resonances are detectable in the corresponding spectra. The native state spectra of F48W and of F22W show the two extreme cases: signals only from tryptophan are observed when both the histidine and tryptophan residues are exposed to the solvent (F48W), and signals only from histidine are observed when the tryptophan side-chain is completely buried in the core (F22W) or absent (WT

HPr). In the native CIDNP spectra of the WT protein (and the F22W mutant), the intensities of the His15 resonances are larger than those of His76 in accord with the greater accessibility of the former. Upon denaturation, however, both histidine residues become fully exposed to the solvent and show similar CIDNP intensities and chemical shifts in the spectrum of the WT protein.

### Photo-CIDNP spectra recorded during refolding

Figure 3 shows CIDNP spectra of the four mutant proteins at two different times after the initiation of refolding by dilution from a high concentration of GdnDCl at 20 °C. Also shown are the spectra of the corresponding native and fully denatured (in 6.0 M GdnDCl) states. After 500 ms, refolding is essentially complete as expected from the refolding rates ranging from 2.86 $s^{-1}$ to 3.80 $s^{-1}$ for the various mutants,[31] and the spectra are almost identical to those of the corresponding native states (except in the case of F22W, see below). On the basis of a cooperative two-state folding model, in which folding proceeds from a fully denatured state directly to the native state, one would expect a series of spectra that correspond to weighted averages of the corresponding spectra of these two states at each time point during folding. In the spectra recorded after 50 ms of folding, e.g. that of F48W, one can indeed observe tryptophan signals corresponding to the fully denatured state and the native state. In addition, however, the spectra show signals with similar intensities that can be attributed to those of one or both of the histidine residues. The CIDNP intensity for a given residue is determined by the result of a competition between the three processes mentioned above:

$$I_{\mathrm{His}} = p_{\mathrm{His}} \frac{k_{\mathrm{His}}[\mathrm{His}]}{k_{\mathrm{q}} + k_{\mathrm{Trp}}[\mathrm{Trp}] + k_{\mathrm{His}}[\mathrm{His}]} \quad (1)$$

$$I_{\mathrm{Trp}} = p_{\mathrm{Trp}} \frac{k_{\mathrm{Trp}}[\mathrm{Trp}]}{k_{\mathrm{q}} + k_{\mathrm{Trp}}[\mathrm{Trp}] + k_{\mathrm{His}}[\mathrm{His}]} \quad (2)$$

where $p_{\mathrm{His}}$ and $p_{\mathrm{Trp}}$ are the polarisations of histidine and tryptophan produced per radical pair in the absence of competition; $k_{\mathrm{His}}$ and $k_{\mathrm{Trp}}$ are the second order rate constants for reaction of histidine and tryptophan with excited flavin respectively, and $k_{\mathrm{q}}$ is the (pseudo-) first order rate constant for the decay of an excited flavin molecule by processes that include quenching by molecular oxygen; and [His] and [Trp] are the concentrations of the exposed side-chains of histidine and tryptophan, respectively. By taking the ratio of these two

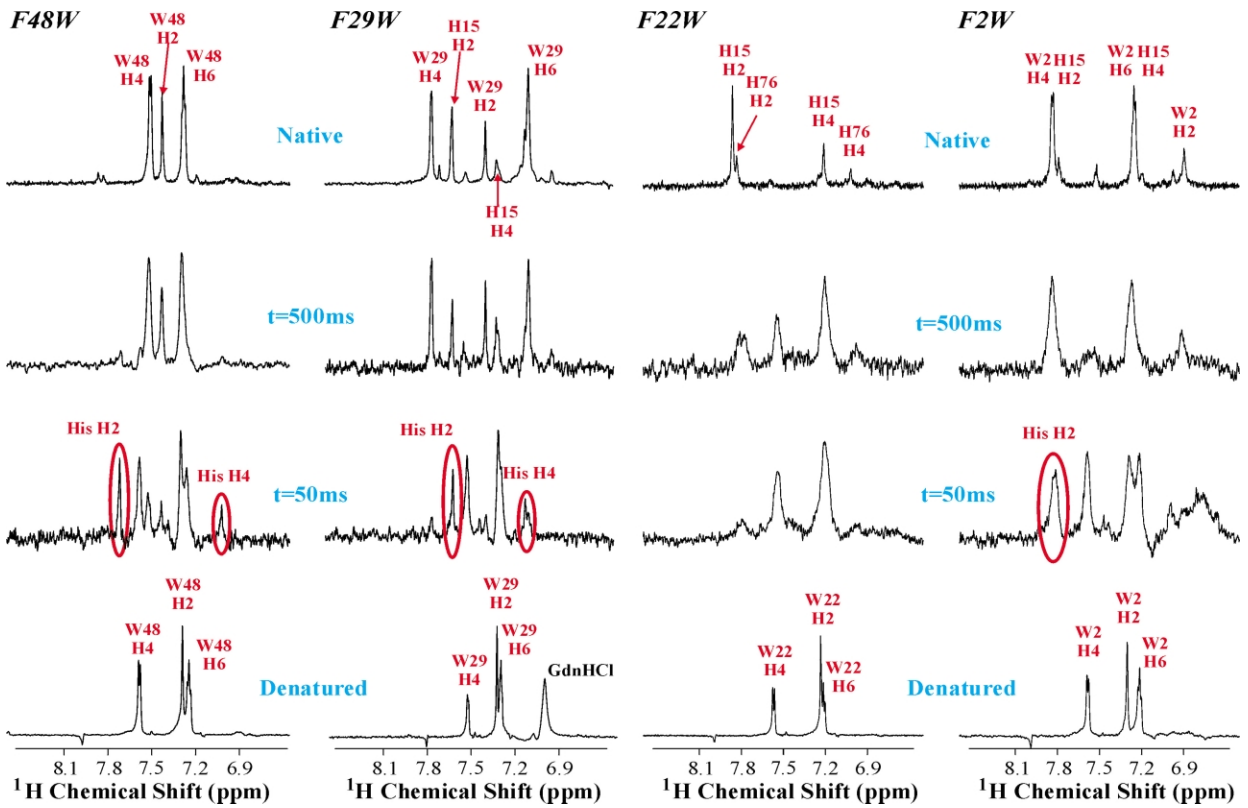


**Figure 3**. Kinetic refolding of the four mutants monitored by photo-CIDNP NMR spectroscopy at 20 °C. The kinetics were followed after a ten-fold dilution into buffer solution at pH 7.0 from 6.0 M GdnDCl. For each mutant, the figure shows spectra of the denatured state (bottom) and native state (top), and kinetic spectra recorded 50 ms and 500 ms after the initiation of refolding.

expressions, one obtains:

$$\frac{I_{His}}{I_{Trp}} = \frac{p_{His}k_{His}[His]}{p_{Trp}k_{Trp}[Trp]} \approx \frac{k_{His}[His]}{k_{Trp}[Trp]} \quad (3)$$

using the fact that the intrinsic polarizations of the two amino acid types are of similar magnitude ($p_{His} \approx p_{Trp}$).[16] When a single-tryptophan mutant of HPr is unfolded in 6 M GdnDCl, the histidine and tryptophan side-chains are likely to be equally accessible (i.e. [His] ≈ 2[Trp]) and hence from equation (3) and using $k_{Trp} : k_{His} \approx 40 : 1$ :[10]

$$\frac{I_{His}}{I_{Trp}} \approx \frac{2k_{His}}{k_{Trp}} \text{ i.e. } I_{His} \ll I_{Trp} \quad (4)$$

The small histidine polarization (~5%) predicted by equation (4) is consistent with the spectra measured for all four denatured states of the mutant proteins (see Figure 2). On the other hand, in the 50 ms refolding spectrum of F48W the intensities of the histidine and tryptophan CIDNP signals are comparable (i.e. $I_{His} \approx I_{Trp}$), which suggests from equation (3):

$$\frac{[Trp]}{[His]} \approx \frac{k_{His}}{k_{Trp}} \text{ or } [Trp] \ll [His] \quad (5)$$

and therefore indicates that a dramatic change in the relative exposures of the histidine and tryptophan residues occurs between the unfolded state and the kinetic intermediate state detected here.

The 50 ms CIDNP spectra recorded during the refolding of F2W and F29W also show additional intense histidine signals, as observed in the case of F48W. The CIDNP spectra recorded during the refolding of F22W, however, show no evidence of the transient histidine peaks seen for the other three mutants. In addition, the rate of appearance of the signals of His15 in the native state protein during refolding is also somewhat slower than in the case of F2W and F29W. These observations may be connected with the absence of an exposed tryptophan side-chain in the native state (all other mutants show at least some tryptophan CIDNP intensity in their native states). Indeed, simulations of the CIDNP kinetics using equations (1) and (2) show that the competition of a rapidly reacting tryptophan in the denatured state with a slowly reacting histidine ($k_{His} = k_{Trp}/40$) in the native state leads to the slow appearance of the native state CIDNP signals (data not shown).

Figure 4 shows the refolding of F48W at low temperature (5 °C) as monitored by photo-CIDNP spectroscopy, and again shows the appearance of intense histidine signals in the earlier time points that gradually disappear with a time dependence similar to that observed for the peaks of the denatured state as folding proceeds to the native state. The quality and time resolution of the results at this temperature, however, enable quantitative information concerning the rates of folding to be extracted. The CIDNP intensity ratio of the β-$CH_2$ signals of the native and intermediate states as a function of the refolding time $t$ was fitted to equation (10) in Materials and Methods (Figure 4(B)). Intensities were estimated by multiplying the widths and heights obtained for each well-resolved peak using a fitting procedure that assumes a Lorentzian line shape.

As illustrated in Figure 4, the error increases dramatically with refolding time, and time points beyond $t = 6$ seconds were therefore not included in the fitting procedure. This procedure yields a value for the time constant of 7 ± 0.6 seconds, a value in good agreement with values obtained for the WT protein by real-time NMR (τ of 8.3 seconds) and ANS fluorescence (τ of 8.4 seconds) at the slightly lower temperature of 2.8 °C.[17] It is interesting to note that by using short 50 ms flashes, and limiting their number to ~16, significant decay of CIDNP intensity due to progressive flavin degradation does not seem to occur in these experiments as significant degradation of the flavin would be observed as a progressive decay in the intensities of the native tryptophan signals in the range of refolding times from ~2 seconds onwards. In fact these signals hardly change at all during this interval, despite repeated light flashes.

## Refolding followed by stopped-flow fluorescence intensity and fluorescence anisotropy measurements

### *Refolding of F48W*

The dramatic change in the relative exposures of the histidine and tryptophan residues between the unfolded state and the kinetic intermediate state of F48W detected in the real-time photo-CIDNP experiments prompted us to study the folding of the F48W mutant in more details by measuring the intensity and anisotropy of the tryptophan fluorescence using stopped-flow methods. Kinetic traces for the refolding of F48W are shown in Figure 5.

The values of the fluorescence intensity and anisotropy for the native state of F48W are in good agreement with steady state measurements.[31] Indeed, the fluorescence anisotropy of the native state of F48W is very low, in accord with the location of this residue on the surface of the protein, projecting into the solvent with essentially unrestricted rotational mobility.

Interestingly, the anisotropy trace of F48W changes during refolding from an initial value of 0.046 to the native state value of 0.038 (Figure 5). This change corresponds to an increase in the mobility of the probe as the folding proceeds, a result that is opposite to that expected if the protein were transforming from a disordered unfolded state towards a more highly structured native state.

To investigate this further, we have calculated the expected fluorescence intensity and fluorescence anisotropy values of the unfolded state under the refolding conditions (see Materials and Methods), and compared these to the values

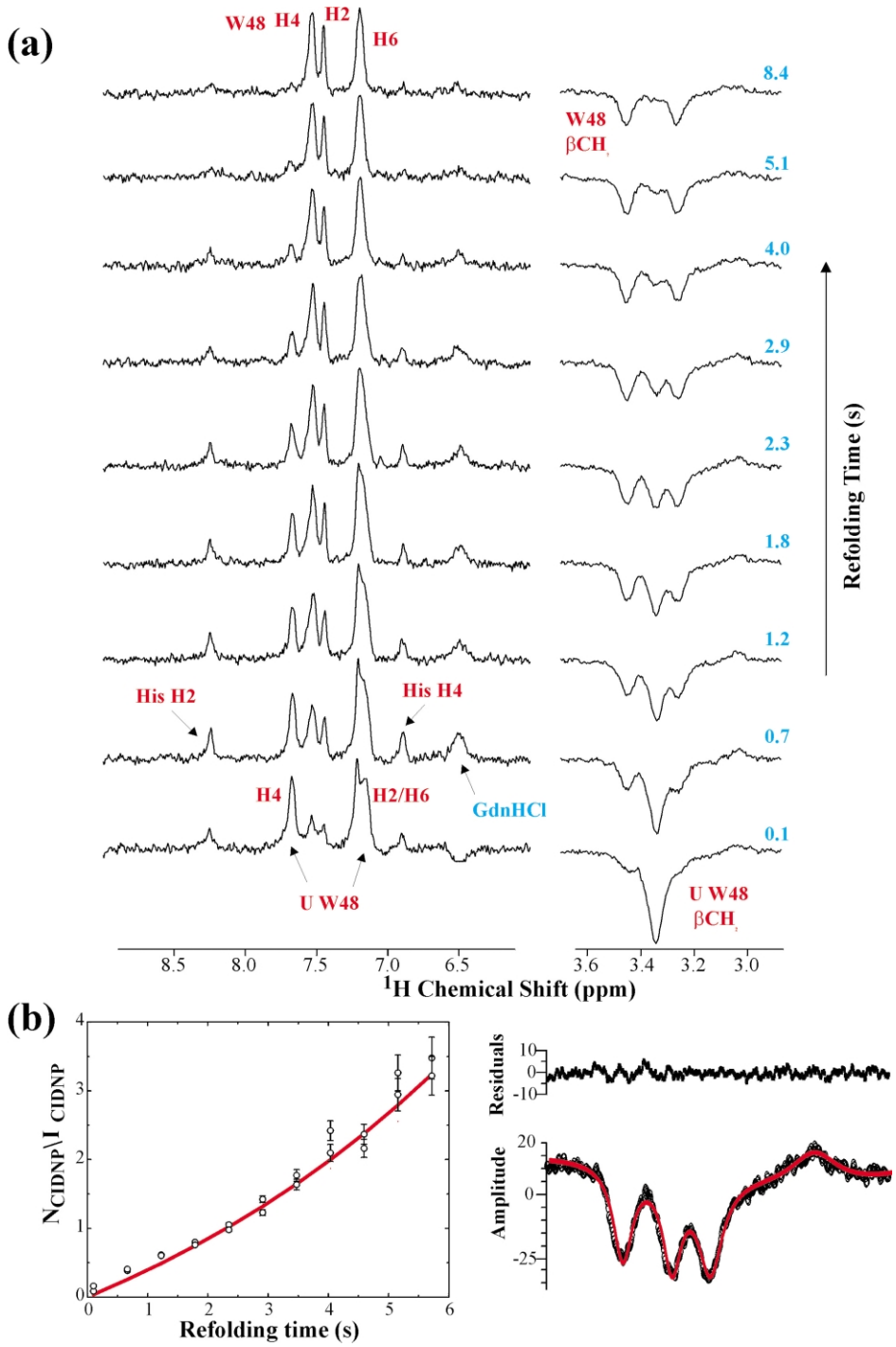


**Figure 4**. Kinetic refolding of the F48W mutant monitored by photo-CIDNP NMR spectroscopy at 5 °C. The kinetics were followed after a ten-fold dilution from 6.0 M GdnDCl. (A) CIDNP spectra acquired at different times after initiation of the refolding reaction. *U* represents the burst-phase intermediate species. The change in the spectra from the intermediate state to the native state is seen particularly clearly in the aliphatic region (right hand side) where the single peak of the $\beta CH_2$ protons of the intermediate state gives way to the two resonances of these protons in the native state. (B) Left panel, ratio of native state to intermediate state CIDNP intensities of the tryptophan $\beta$-$CH_2$ signals shown in (A) as a function of refolding time. The data were fitted as described in equation (10). Right panel: Illustration of the Lorentzian fitting procedure for the $t = 2.3$ seconds spectrum shown in (A).

extrapolated to the initiation of the refolding process. The results provide clear evidence that substantial changes occur in the dead-time of the refolding experiment. The fluorescence intensity of the state formed immediately after the initiation of refolding is significantly higher than that of the unfolded state, with an increase corresponding to 21% of the total observed kinetic change during the entire refolding process. Besides, the fluorescence anisotropy of the fully unfolded state

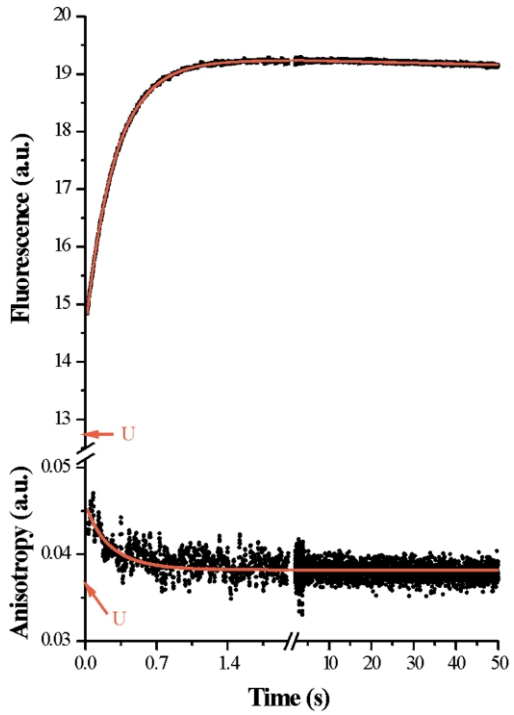


**Figure 5**. Stopped-flow fluorescence and fluorescence anisotropy traces of the refolding of F48W of HPr in 0.54 M GdnHCl at 20 °C and pH 7.0. The two kinetic traces were fitted simultaneously to a reaction scheme involving two states as described previously.[31] The predicted fluorescence intensity and fluorescence anisotropy values of the unfolded state calculated for the refolding conditions are indicated by arrows (see Materials and Methods).

(indicated by an arrow in Figure 5) is lower than that of the native state. It has also been found that the anisotropy decreases during the unfolding of F48W (data not shown), as it does during the refolding.

Taken together, these results therefore indicate that the earliest fluorescence and anisotropy values acquired in the stopped-flow experiment are those of a burst-phase intermediate, which has formed from the unfolded state within the dead-time of the experiment. The anisotropy measurements reveal that the mobility of the tryptophan residue at the position 48 is more restricted in this intermediate than in the native state. These observations are consistent with the partial burial of the side-chain of tryptophan residue 48 during the initial stages of folding observed in the photo-CIDNP experiments.

### *Refolding of the other single tryptophan mutants*

The results obtained by stopped-flow fluorescence anisotropy during the kinetic refolding of the F48W mutants prompted us to investigate the refolding of the other single tryptophan mutants by the same technique. The results are illustrated in Figure 6.

The time dependence of the anisotropy values for the different mutants provides important information about the formation of native-like states during folding. Thus, although the native states all have different values of their anisotropy, reflecting the different positions of the four residues in the structure,[31] the refolding traces reveal that all tryptophan residues display similar anisotropy levels directly after the initiation of the refolding (Figure 6). This shows that the motion of each tryptophan residue in the mutant proteins is restricted to a similar degree after a few ms, suggesting that there is low structural specificity of this burst phase intermediate. It is therefore likely that this intermediate resembles a collapsed state where the various hydrophobic tryptophan residues are buried to a significant extent from the solvent, before any persistent secondary structure is formed.

## Fluorescence quenching confirms the existence of a burst phase collapsed state

Fluorescence quenching has been monitored during the folding of a large number of proteins, such as hen lysozyme[18] and TEM-1 β-lactamase.[19] These experiments, like the photo-CIDNP measurements, probe the degree of exposure to solvent of tyrosine and tryptophan side-chains. In the case of the single-tryptophan mutants of HPr, we have the opportunity to obtain specific local information by monitoring the extent of fluorescence quenching of individual residues during refolding. Figure 7(A) shows the fluorescence signal of F48W during refolding in the presence of increasing amounts of sodium iodide, between 0 mM and 300 mM, at constant ionic strength. The presence of iodide ions causes the native state fluorescence of F48W to be very efficiently quenched, consistent with the exposed location of the probe in the structure (see Figure 1). The burst-phase intermediate is quenched to a lesser extent, as can be seen from the change from a initial increase in fluorescence to a initial decrease in fluorescence as the iodide concentration is raised.

In order to measure the extent of quenching in these different states, a Stern–Volmer plot was constructed, Figure 7(B). Using a simple collision model that predicts a linear relationship between the quencher concentration, [Q], and the reciprocal of the fluorescence intensity, one obtains $F_0/F = 1 + K_{sv}[Q]$, where $F$ and $F_0$ are the intensities in the presence and absence of quencher, respectively, and $K_{sv}$ is the effective quenching constant, the ratio of the rate constant of collisional quenching to the unquenched fluorescence rate constant.[18] The burst-phase intermediate shows a smaller quenching constant ($K_{sv} = 7.7\ M^{-1}$) than does the native state ($K_{sv} = 12.8\ M^{-1}$) corresponding to an increase in tryptophan accessibility as the intermediate converts into native state. This conclusion

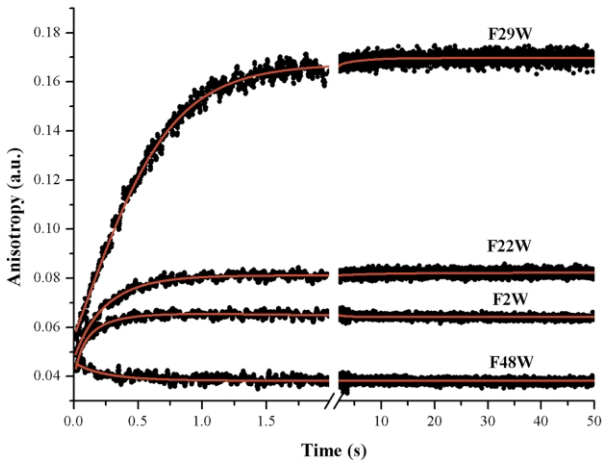


**Figure 6**. Stopped-flow fluorescence anisotropy traces of the refolding of the four mutant proteins in 0.54 M GdnHCl at 20 °C and pH 7.0. The kinetic traces were fitted to a reaction scheme involving two (F48W) or three (F2W, F22W, F29W) states as described previously.[31]

is in agreement with the exclusion from solvent of Trp48 in the initial stage of folding as observed in the photo-CIDNP experiments, and the lower mobility of this state derived from the fluorescence anisotropy measurements.

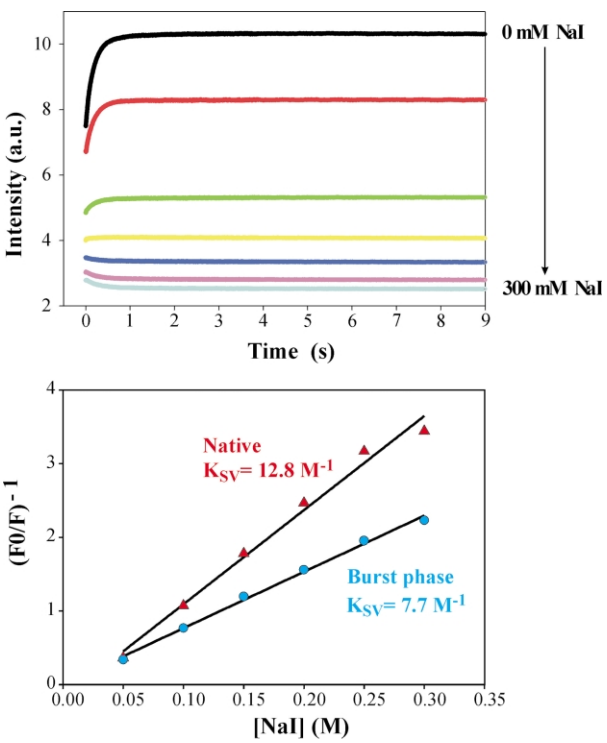


**Figure 7**. Quenching of fluorescence by sodium iodide during the refolding of the F48W mutant of HPr in 0.54 M GdnHCl at 20 °C and pH 7.0. (A) Kinetic traces of the refolding at different NaI concentrations (0, 50, 100, 150, 200, 250 and 300 mM). (B) Stern–Volmer quenching plots calculated from the data in (A).

## Discussion

The combination of real-time NMR measurements with photo-CIDNP methods represents a powerful approach for characterizing directly structural changes taking place in a protein during folding. The photo-CIDNP results presented here indicate that after the initiation of the refolding reaction, HPr collapses around at least three (2, 29, and 48) out of four tryptophan side-chains, making them less accessible to the flavin dye to an extent that the histidine residues at position 15 and/or 76 can compete effectively for the dye and become polarized. The behaviour of F48W is particularly interesting and is illustrated schematically in Figure 8. In both the fully denatured and native states, the tryptophan side-chain is exposed and can react sufficiently rapidly with the excited state of the flavin dye that the histidine residues acquire negligible polarisation even if both these residues are also exposed. In the species formed once refolding is initiated, however, the tryptophan residue is partially buried, allowing one or both histidine residues to react with the flavin dye and become polarized. Iodide fluorescence quenching experiments with the F48W mutant (Figure 7) show that the quenching is less efficient for the intermediate state formed in the dead-time of the stopped-flow experiment than for the native state. These results are consistent with the formation of a collapsed compact denatured state put forward to explain the photo-CIDNP data. Further evidence for a collapse of structure around the tryptophan side-chain comes from fluorescence anisotropy measurements (Figure 5). These data show that the tryptophan side-chain of the F48W mutant has lower mobility in the burst phase species than in both denatured and native states, a result consistent with the observations from the photo-CIDNP and fluorescence quenching studies.

The very rapid hydrophobic collapse observed within a few ms after the initiation of refolding is likely to occur because of favourable free energy

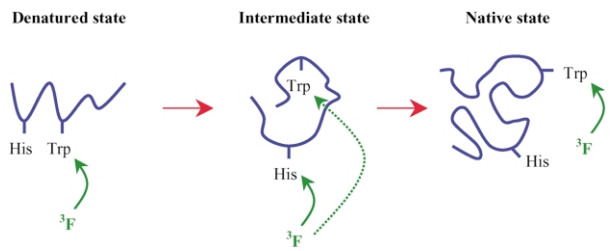


**Figure 8**. Diagram summarizing the observed photo-CIDNP behaviour for F48W. Only one histidine residue is shown for simplicity. In both denatured (left) and native (right) states, the tryptophan side-chain is exposed and reacts so efficiently with the flavin dye that the histidine residue is not measurably polarized even though exposed. In the collapsed state (middle) the tryptophan side-chain is partially buried, allowing the histidine residues to react and become polarized.

associated with the burial of hydrophobic residues once the concentration of GdnHCl is reduced. It is interesting, however, that the burial of some residues at early stages of the folding process, even though it involves non-native contacts, could be important to reduce the propensity of the protein to aggregate by reducing the exposed hydrophobic surface area. In the native state, the energetics of side-chain packing allows the exposure of a number of specific hydrophobic residues for functional purposes. Indeed, the exposed phenylalanine residue at position 48 in HPr is involved in binding to the cytoplasmic and membrane-bound enzymes (enzyme I and II, respectively) to mediate the transfer of a phosphoryl group between the two enzymes.[20,21]

The photo-CIDNP data presented here are similar to the observation for lysozyme where functionally important tryptophan residues are buried initially but emerge when the native state is formed.[7] This observation, and the related finding that non-native interactions are present in denatured lysozyme, involving the active site residue Trp62,[22] have also been explained in terms of the burial during the initial stages of folding of hydrophobic residues, which are functionally important in the native state, in order to prevent aggregation. Radford and co-workers have recently observed similar events during the folding of a small helical protein, Im7, where it was found that two hydrophobic leucine residues initially get buried in the intermediate state, and later become exposed to the solvent in the native state, where they play a key role in binding to the endonuclease domain of the bacterial toxin colicin E7.[23] The results presented here therefore provides insight into a phenomenon that could be rather general, and support the idea that the sequence of a protein codes for a range of characteristics of other than those that are required simply to stabilize the native fold.

## Materials and Methods

### Protein production

The production, purification and characterization of the four single-tryptophan mutants F2W, F22W, F29W, and F48W and the WT HPr were carried out as described elsewhere.[31]

### Photo-CIDNP methods

#### *Equilibrium photo-CIDNP experiments*

Spectra were obtained essentially as described previously.[7,24,25] Blue–green light (4 W) from a continuous-wave argon ion laser (Spectra Physics Stabilite 2016-05) was chopped into 100 ms pulses by a mechanical shutter controlled from a home-built 600 MHz spectrometer located at the Oxford Center for Molecular Sciences. The light was introduced into the NMR tube *via* an optical fiber (diameter 1 mm), the end of which was held 4 mm above the top of the NMR coil, inside a coaxial insert (Wilmad WGS 5BL) dipping into the solution.[26] Flavin mononucleotide (0.2 mM) was used as photosensitiser in all experiments.

#### *Real-time photo-CIDNP experiments*

The rapid injection device used for all real-time photo-CIDNP experiments was similar to that described earlier.[6,7,27] To enable rapid transfer of denatured protein into the NMR tube, a thin (0.5 mm internal diameter) PTFE tube was introduced into the sample via a small hole in the coaxial insert and connected to a pneumatically driven syringe outside the magnet. The syringe was embedded in a solid PTFE block to absorb mechanical shock and was mounted on an aluminium support along with the pneumatic triggering system. The syringe and PTFE transfer line were washed several times with $^2H_2O$, and a small amount (50 μl) of denatured protein was taken up into the end of the transfer line. The NMR tube was then filled with refolding buffer solution (500 μl) and the coaxial insert with optical fiber and PTFE tubing was carefully placed into the NMR tube until both the insert and tubing dipped below the surface of the buffer solution, 4 mm above the top of the NMR coil. With this arrangement, no significant loss in field homogeneity was observed and adequate line widths were easily obtained. Shimming and tuning were first performed on a separate sample with the denatured protein injected into the refolding buffer and then left unchanged for further experiments.

In each refolding experiment a 10 bar, 30 ms injection pulse was used to initiate folding. After an initial recovery delay (i.e. greater than or equal to 50 ms) a hard 90° radiofrequency pulse and field gradient pulse were applied to destroy all background magnetization. A 50 ms laser flash was then applied to generate polarisation in exposed aromatic side-chains and a CIDNP spectrum was acquired. When folding was fast (e.g. at 20 °C, where $k_{obs} \sim 3\,s^{-1}$) the entire process was repeated with a new sample and a different initial delay to obtain the time-dependence of the spectrum. When folding was sufficiently slow (e.g. at 5 °C, where $k_{obs} \sim 0.1\,s^{-1}$) a series of CIDNP spectra were acquired following a single injection with the delay between CIDNP spectra limited by the spectral acquisition time (>250 ms). All spectra were the result of the acquisition a single free induction decay (FID).

The quality and time resolution of the results at low temperature enabled quantitative information concerning the rates of folding to be extracted. After the initial formation of the intermediate state, folding proceeds in a cooperative two-state process to the final native state. The system can be treated as a binary mixture in which the concentrations of intermediate (I) and native (N) states as a function of refolding time $t$ are given, respectively, by:

$$[\mathrm{I}] \propto e^{-t/\tau} \text{ and } [\mathrm{N}] \propto 1 - e^{-t/\tau} \quad (6)$$

where $\tau$ is the folding time constant. The competition for the flavin dye at any time will therefore be between exposed residues in both the intermediate and native states, and between these and any other existing quenching mechanism. If an arbitrary number $n$ of potentially CIDNP active residues are present in the protein, the CIDNP intensity $C_l^I$ for a given residue l in

the intermediate state can be represented by:

$$C_{\mathrm{i}}^{\mathrm{I}} \propto \frac{k_1^{\mathrm{I}}[\mathrm{I}]}{k_{\mathrm{q}} + \sum_{i=1}^{n} k_i^{\mathrm{I}}[\mathrm{I}] + \sum_{i=1}^{n} k_i^{\mathrm{N}}[\mathrm{N}]} \quad (7)$$

where $k_{\mathrm{q}}$ is the (pseudo-) first order rate constant for the decay of triplet flavin by fluorescence and/or quenching by molecular oxygen and $k_i^{\mathrm{N}}$ and $k_i^{\mathrm{I}}$ are the second order rate constants for quenching of triplet flavin by residue i in the native and intermediate state, respectively. Similarly, the CIDNP intensity for residue j in the native state will be given by:

$$C_{\mathrm{j}}^{\mathrm{N}} \propto \frac{k_1^{\mathrm{N}}[\mathrm{N}]}{k_{\mathrm{q}} + \sum_{i=1}^{n} k_i^{\mathrm{I}}[\mathrm{I}] + \sum_{i=1}^{n} k_i^{\mathrm{N}}[\mathrm{N}]} \quad (8)$$

These two equations illustrate the complexity introduced into the interpretation of CIDNP intensities by competition for the excited flavin molecules. The situation can be simplified somewhat by taking the ratio of equations (8) and (7) to give:

$$\frac{C_{\mathrm{j}}^{\mathrm{N}}}{C_{\mathrm{i}}^{\mathrm{I}}} \propto \frac{[\mathrm{N}]}{[\mathrm{I}]} \quad (9)$$

and finally, by inserting expressions from equation (6) and simplifying, one obtains:

$$\frac{C_{\mathrm{j}}^{\mathrm{N}}}{C_{\mathrm{i}}^{\mathrm{I}}} \propto \mathrm{e}^{t/\tau} - 1 \quad (10)$$

By plotting the CIDNP intensity ratio of any peak in the native state and any peak in the intermediate state against the refolding time $t$ and fitting the data with equation (10) one can extract a value for the folding time constant $\tau$.

### Stopped-flow fluorescence, fluorescence quenching and anisotropy measurements

Stopped-flow fluorescence experiments were carried out using a BioLogic SFM3 mixer (BioLogic, Claix, France) with a 2 mm × 2 mm cell (BioLogic FC-20) and a BioLogic PMS 400 detection system. The experimental details have been described previously.[31] The fluorescence intensity and fluorescence anisotropy levels of the unfolded F48W mutant protein in the final refolding conditions were calculated by taking measurements in 4, 5 and 6 M GdnHCl, and extrapolating these to the final refolding conditions (0.54 M GdnHCl). This method accounts for the effect of GdnHCl concentration on the fluorescence intensity in the absence of conformational changes,[18,28] and for the effect of GdnHCl concentration on fluorescence anisotropy that are likely the result from changes in the viscosity of the solutions.[29]

## Acknowledgements

We thank Dr K. Maeda (Oxford University) for providing us with the Lorentzian fitting algorithm used in this study. This work has been supported by the INTAS project (01-2126), and by the TMR Network "Structure and dynamics of intermediate states in protein folding" (FMRX960013) and RTD project "Transient NMR" (HPRI-CT-1999-50006) of the European Union. The Oxford Centre for Molecular Sciences is supported by EPSRC, BBSRC and MRC. The research of C.M.D. is also supported by a Programme Grant from the Wellcome Trust.

## References

1. Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332.
2. Radford, S. E. & Dobson, C. M. (1999). From computer simulations to human disease: emerging themes in protein folding. *Cell*, **97**, 291–298.
3. Dobson, C. M. (2002). Getting out of shape. *Nature*, **418**, 729–730.
4. Fersht, A. R. & Daggett, V. (2002). Protein folding and unfolding at atomic resolution. *Cell*, **108**, 573–582.
5. Ellis, R. J. & Pinheiro, T. J. T. (2002). Medicine: danger—misfolding proteins. *Nature*, **416**, 483–484.
6. Van Nuland, N. A. J., Balbach, J., Forge, V. & Dobson, C. M. (1998). Real-time NMR studies of protein folding. *Accts. Chem. Res.* **31**, 773–780.
7. Hore, P. J., Winder, S. L., Roberts, C. H. & Dobson, C. M. (1997). Stopped-flow photo-CIDNP observation of protein folding. *J. Am. Chem. Soc.* **119**, 5049–5050.
8. Maeda, K., Lyon, C. E., Lopez, J. J., Cemazar, M., Dobson, C. M. & Hore, P. J. (2000). Improved photo-CIDNP methods for studying protein structure and folding. *J. Biomol. NMR*, **16**, 235–244.
9. Wirmer, J., Kühn, T. & Schwalbe, H. (2001). Millisecond time resolved Photo-CIDNP NMR reveals a non-native folding intermediate on the ion induced refolding pathway of bovine-Lactalbumin. *Angew. Chem.* **113**, 4378–4381.
10. Winder, S. L., Broadhurst, R. W. & Hore, P. J. (1995). Photo-CIDNP of amino acids and proteins: effects of competition for flavin triplets. *Spectrochim. Acta part A*, **51**, 1753–1761.
11. Van Nuland, N. A. J., Hangyi, I. W., Van Schaik, R. C., Berendsen, H. J. C., Van Gunsteren, W. F., Scheek, R. M. & Robillard, G. T. (1994). The high-resolution structure of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from nuclear magnetic resonance nuclear Overhauser effect data. *J. Mol. Biol.* **237**, 544–559.
12. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1993). Phosphoenol-pyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–594.
13. Muszkat, K. A. & Wismontski-Knittel, T. (1985). Reactivities of tyrosine, histidine, tryptophan, and methionine in radical pair formation in flavin triplet induced protein nuclear magnetic polarization. *Biochemistry*, **24**, 5416–5421.
14. Tsentalovich, Yu. P., Lopez, J. J., Hore, P. J. & Sagdeev, R. Z. (2002). Mechanisms of reactions of flavin mononucleotide triplet with aromatic amino acids. *Spectrochim. Acta part A*, **58**, 2043–2050.
15. Morozova, O. B., Yurkovskaya, A. V., Tsentalovich, Yu. P., Forbes, M. D. E., Hore, P. J. & Sagdeev, R. Z. (2002). Time resolved CIDNP study of electron transfer reactions in proteins and model compounds. *Mol. Phys.* **100**, 1187–1195.

16. Hore, P. J. & Broadhurst, R. W. (1993). Photo-CIDNP of biopolymers. *Prog. Nuci. Magn. Reson. Spectrosc.* **25**, 345–402.
17. Van Nuland, N. A. J., Meijberg, W., Warner, J., Forge, V., Scheek, R. M., Robillard, G. T. & Dobson, C. M. (1998). Slow cooperative folding of a small globular protein HPr. *Biochemistry*, **37**, 622–637.
18. Itzhaki, L. S., Evans, P. A., Dobson, C. M. & Radford, S. E. (1994). Tertiary interactions in the folding pathway of hen lysozyme: kinetic studies using fluorescent probes. *Biochemistry*, **33**, 5212–5220.
19. Vanhove, M., Lejeune, A., Guillaume, G., Virden, R., Pain, R. H., Schmid, F. X. & Frere, J. M. (1998). A collapsed intermediate with nonnative packing of hydrophobic residues in the folding of TEM-1 beta-lactamase. *Biochemistry*, **37**, 1941–1950.
20. Van Nuland, N. A. J., Kroon, G. J. A., Dijkstra, K., Wolters, G. K., Scheek, R. M. & Robillard, G. T. (1993). The NMR determination of the IIA(mtl) binding site on HPr of the *Escherichia coli* phosphoenol pyruvate-dependent phosphotransferase system. *FEBS Letters*, **315**, 11–15.
21. Van Nuland, N. A. J., Boelens, R., Scheek, R. M. & Robillard, G. T. (1995). High-resolution structure of the phosphorylated form of the histidine-containing phosphocarrier protein HPr from *Escherichia coli* determined by restrained molecular dynamics from NMR-NOE data. *J. Mol. Biol.* **246**, 180–193.
22. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T. *et al.* (2002). Long-range interactions within a non-native protein. *Science*, **295**, 1719–1722.
23. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nature Struct. Biol.* **3**, 209–216.
24. Kaptein, R. (1978). *NMR Spectroscopy in Molecular Biology* (Pullman, B., ed.), pp. 211–229, D. Reidel, Dordrecht.
25. Kaptein, R., Dijkstra, K. & Nicolay, K. (1978). Laser photo-CIDNP as a surface probe for proteins in solution. *Nature*, **274**, 293–294.
26. Scheffler, J. E., Cottrell, C. E. & Berliner, L. J. (1985). An inexpensive, versatile sample illuminator for photo-CIDNP on any NMR spectrometer. *J. Magn. Reson.* **63**, 199–201.
27. Balbach, J., Forge, V., van Nuland, N. A. J., Winder, S. L., Hore, P. J. & Dobson, C. M. (1995). Following protein folding in real time using NMR spectroscopy. *Nature Struct. Biol.* **2**, 865–870.
28. Matagne, A., Chung, E. W., Ball, L. J., Radford, S. E., Robinson, C. V. & Dobson, C. M. (1998). The origin of the alpha-domain intermediate in the folding of hen lysozyme. *J. Mol. Biol.* **277**, 997–1005.
29. Canet, D., Doering, K., Dobson, C. M. & Dupont, Y. (2001). High-sensitivity fluorescence anisotropy detection of protein-folding events: application to alpha-lactalbumin. *Biophys. J.* **80**, 1996–2003.
30. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
31. Azuaga, A. I., Canet, D., Smeenk, G., Berends, R., Titgemeyer, F., Duurkens, R. *et al.* (2003). Characterization of single-tryptophan mutants of Histidine-containing Phosphocarrier Protein: Evidence for local rearrangements during folding from high concentrations of denaturant. *Biochemistry*, **42**, 4883–4895.

***Edited by P. Wright***

*(Received 9 January 2003; received in revised form 31 March 2003; accepted 7 April 2003)*